On the growth of components with non fixed excesses

Anne-Elisabeth Baert, Vlady Ravelomanana, Loÿs Thimonier

On the growth of components with non fixed excesses

Anne-Elisabeth Baert [a], Vlady Ravelomanana [b] and Loys Thimonier [a]

[a] *LaRIA CNRS EA 2083, Université de Picardie Jules-Verne*
33, Rue du Moulin-Neuf
80000 Amiens, France

[b] *LIPN CNRS UMR 7030, Université de Paris-Nord.*
99, av. J.B. Clement
93430 Villetaneuse, France

Abstract

Denote by an l-component a connected graph with l edges more than vertices. We prove that the expected number of creations of $(l+1)$-component, by means of adding a new edge to an l-component in a randomly growing graph with n vertices, tends to 1 as l, n tends to ∞ but with $l = o(n^{1/4})$. We also show, under the same conditions on l and n, that the expected number of vertices that ever belong to an l-component is $\sim (12l)^{1/3} n^{2/3}$.

Key words: Random graphs; asymptotic enumeration; Wright's coefficients.

1 Introduction

We consider here simple labelled graphs, i.e., graphs with labelled vertices, without self-loops and multiple edges. A random graph is a pair $(\mathcal{G}, P)$ where $\mathcal{G}$ is a family of graphs and P is a probability distribution over $\mathcal{G}$. Topics on random graphs provide a large and particularly active body of research. For excellent books on these fields, see Bollobas [3] or Janson [8]. In this paper, we consider the *continuous time* random graph model $\{\mathbb{G}(n,t)\}_t$ which consists on assigning a random variable, T_e, to each edge e of the complete graph K_n. The $\binom{n}{2}$ variables T_e are independent with a common continuous distribution and the edge set of $\{\mathbb{G}(n,t)\}_t$ is constructed with all edge e such that $T_e \leq t$. Throughout this paper, a $(k, k+l)$ graph is one having k vertices and $k+l$ edges. its *excess* is k. A $(k, k+l)$ *connected* graph is said an *l-component.*

These terms are due, as far as we know, respectively to Janson in [6] and to Janson *et al* in [7].

To obtain the results presented here, methods of the probabilistic model $\{\mathbb{G}(n,t)\}_t$, studied in [6], are combined with asymptotic enumeration methods (having the "counting flavour") developped by Wright in [10] and by Bender *et al.* in [1,2]. The problems we consider are in essence combinatorial problems. Often, combinatorics and probability theory are closely related and the approaches given here furnish efficient uses of methods of asymptotic analysis to get extreme characteristics of random labelled graphs.

Following Janson in [6], we define by $\alpha(l;k)$, the expected number of times that a new edge is added to an l-component of order k which becomes an $(l+1)$-component. This transition is denoted $l \to l+1$. It has been proved by Janson *et al.* in [7] that with probability tending to 1, there is exactly one such transition, see Theorem 16 and its proof. Our purpose in this paper is to give a different approach to obtain this result, when $l \to \infty$ with n. More precisely, we propose an alternative method and direct calculations of the difficult results given in [7], connecting these results to those of Bender *et al.* in [1,2]. To do this, we evaluate $\alpha_l = \sum_{k=1}^{n} \alpha(l;k)$, the expected number of transitions $l \to l+1$, and show that, whenever $l \to \infty$ with n but $l = o(n^{1/4})$, $\alpha_l \sim 1$.

Moreover, let V_l be the number of vertices that ever belong to an l-component and V_l^{max} the order of the largest l-component that ever appears. We prove that, whenever $l \equiv l(n) \to \infty$ with n but $l = o(n^{1/4})$, $V_l \sim (12l)^{1/3} n^{2/3}$ and $V_l^{max} = O(l^{1/3} n^{2/3})$. In particular, these results improve the work of Janson and confirm his predictions (cf. [6, Remark 8]).

2 The Expected Number of Creations of $(l+1)$-excess Graphs

When adding an edge in a randomly growing graph, there is a possibility that it joins two vertices of the same component, increasing its excess. Let $\alpha(l;k)$ be the expected number of times that a new edge is added to an l-component of order k. Denote by $c(k,k+l)$ the number of connected $(k,k+l)$-graph then we have the following lemma.

Lemma 1 *For all $l \geq -1$ and $k \geq 1$, we have*

$$\alpha(l;k) = (n)_k \frac{(k+l)!}{k!} c(k,k+l) \frac{(k^2-3k-2l)}{2} \frac{(nk - k^2/2 - 3k/2 - l - 1)!}{(nk - k^2/2 - k/2)!} \tag{1}$$

and for all l, k and n such that $l = O(k^{2/3})$, $1 \leq k \leq n$

$$\alpha(l;k) = \frac{1}{2}\rho_l \frac{k^{(3l+1)/2}}{n^{l+1}} \exp\left(-\frac{k^3}{24n^2} + \frac{lk^2}{8n^2} + \frac{lk}{2n}\right)$$
$$\times \left(1 + O\left(\frac{k}{n} + \frac{k^4}{n^3} + \frac{1}{k}\right) + O\left(\frac{l^3}{k^2} + \frac{l^{1/2}}{k^{1/2}} + \frac{(l+1)^{1/16}}{k^{9/50}}\right)\right) \tag{2}$$

where

$$\rho_l = \frac{1}{2}\sqrt{\frac{3}{\pi}}\left(\frac{e}{12l}\right)^{\frac{l}{2}}(1 + O(1/l)) \tag{3}$$

◇

Before proving lemma 1, let us recall the extension, due to Bender, Canfield and McKay [2], of Wright's formula for $c(k, k+l)$, the asymptotic number of connected sparsely edged graphs [10].

Theorem 2 (Bender-Canfield-McKay 1992) *There exists sequence r_i of constants such that for each fixed $\epsilon > 0$ and integer $m > 1/\epsilon$, the number $c(k, k+l)$ of connected sparsely edged graphs satisfies*

$$c(k, k+l) = \sqrt{\frac{3}{\pi}}\frac{w_l}{2}\left(\frac{e}{12l}\right)^{l/2} k^{k+(3l-1)/2}$$
$$\times \exp\left(\sum_{i=1}^{m-2} \frac{r_i l^{i+1}}{k^i} + O\left(\frac{l^m}{k^{m-1}} + \sqrt{\frac{l}{k}} + \frac{(l+1)^{1/16}}{k^{9/50}}\right)\right) \tag{4}$$

uniformly for $l = O(k^{1-\epsilon})$. The first few values of the constant r_i are

$$r_1 = -\frac{1}{2},\ r_2 = \frac{701}{2100},\ r_3 = -\frac{263}{1050},\ r_4 = \frac{538\,859}{2\,695\,000}.$$

◇

Note that, the factor w_l in (4) is given, for $l > 0$, by

$$w_l = \pi \frac{\Gamma(l)}{\Gamma(3l/2)} d_l \sqrt{\frac{8}{3}}\left(\frac{27l}{8e}\right)^{l/2} \tag{5}$$

where $d_l = 1/(2\pi) + O(1/l)$ and $w_0 = \pi/\sqrt{6}$ (see [1,10]). We also remark here that in lemma 1, we restrict our attention to values of l such that $l = O(k^{2/3})$ which will be shown to be sufficient to obtain the result in theorem 4.

Proof of lemma 1. The proof given here are based on the works of Janson in [5,6]. However, the main difference comes from the fact that our parameter, representing the excess of the sparse components l, is no more fixed as in [6]. When a new edge is added to an l-component of order k, there are $\binom{n}{k}c(k, k+l)$ manners to choose an l-component and $\binom{k}{2} - k - l$ ways to choose the new edge. Furthermore, the probability that such possible component is one of $\{\mathbb{G}(n,t)\}_t$ is $t^{k+l}(1-t)^{(n-k)k+\binom{k}{2}-k-l}$ and with the conditional probability $\frac{dt}{(1-t)}$ that a given edge is added during the interval $(t, t+dt)$ and not earlier, integrating over all times, we obtain

$$\alpha(l;k) = \binom{n}{k} c(k,k+l) \left(\frac{k^2-3k-2l}{2} \right) \int_0^1 t^{k+l}(1-t)^{(n-k)k+\binom{k}{2}-k-l-1} dt \quad (6)$$

which evaluation leads to (1). For $1 \le k \le n$ and $l = O(k^{2/3})$, the value of the integral in (6) is

$$\frac{(nk-k^2/2-3k/2-l-1)!}{(nk-k^2/2-k/2)!} = k^{-k-l-1}(n-k/2)^{-k-l-1} \left(1 + O(\frac{k}{n}) \right) . \quad (7)$$

Furthermore,

$$\frac{(n)_k}{(n-k/2)^k} = \exp\left(-\frac{k^3}{24n^2} \right) \left(1 + O(k/n + k^4/n^3) \right) \quad (8)$$

and obviously

$$\binom{k}{2} - k - l = \frac{k^2}{2}\left(1 + O(1/k)\right) . \quad (9)$$

Thus, combining (7),(8) and (9) in (1), we infer that

$$\alpha(l;k) = \frac{1}{2}\frac{(k+l)!}{k!} c(k,k+l) \frac{\exp\left(-\frac{k^3}{24n^2}\right)}{(n-k/2)^{l+1}k^{k+l-1}} \left(1 + O(1/k + k/n + k^4/n^3) \right) . \quad (10)$$

Using Taylor expansions

$$\ln\left(\frac{(k+l)!}{k^l k!} \right) = \frac{l^2}{2k} + O(l^3/k^2) + O(l/k) \quad (11)$$

which is sufficient for our present purpose ($l = O(k^{2/3})$). Also, we get

$$\frac{(n-k/2)^{l+1}}{n^{l+1}} = \exp\left(-\frac{lk}{2n} - \frac{lk^2}{8n^2} \right) \left(1 + O(k/n + lk^3/n^3) \right) . \quad (12)$$

Note that $\frac{k^4}{n^3}$ dominates the term $\frac{lk^3}{n^3}$ in (12). Then, using the asymptotic formula for the number $c(k, k+l)$ given by theorem 2 in (10), we obtain (2). $\square$

The form given by equation (2), in lemma 1, suggests us to consider the asymptotic behaviour of

$$\sum_{k=1}^{n} k^a \exp\left(-\frac{k^3}{24n^2} + \frac{lk^2}{8n^2} + \frac{lk}{2n} \right)$$

where $a = \frac{3l+1}{2}$, $l \equiv l(n)$ as $n \to \infty$.

Lemma 3 *As $n \to \infty$, $l \equiv l(n) = o(n^{1/4})$ and $a = \frac{3l+1}{2}$, we have*

$$\sum_{k=1}^{n} k^a \exp\left(-\frac{k^3}{24n^2} + \frac{lk^2}{8n^2} + \frac{lk}{2n}\right) \sim 2^{a+1} 3^{(a-2)/3} \Gamma\left(\frac{a+1}{3}\right) n^{2(a+1)/3} . \tag{13}$$

$\diamond$

Proof. We start estimating the summation by an integral using, for e.g., the classical Euler-Maclaurin method for asymptotics estimates of summations (see [4]),

$$\sum_{k=1}^{n} k^a exp\left(-\frac{k^3}{24n^2} + \frac{lk^2}{8n^2} + \frac{lk}{2n}\right) \sim \int_0^n t^a \exp\left(-\frac{t^3}{24n^2} + \frac{lt^2}{8n^2} + \frac{lt}{2n}\right) dt \tag{14}$$

If we denote by I_n the integral, we have after substituting $t = 2n^{2/3}e^z$:

$$I_n \sim 2^{a+1} n^{\frac{2(a+1)}{3}} \int_{-\infty}^{+\infty} \exp(h(z))dz \tag{15}$$

where

$$h(z) = -\frac{1}{3}e^{3z} + \frac{l}{2n^{2/3}}e^{2z} + \frac{l}{n^{1/3}}e^z + (a+1)z . \tag{16}$$

We have

$$h'(z) = -e^{3z} + \frac{l}{n^{2/3}}e^{2z} + \frac{l}{n^{1/3}}e^z + (a+1) \tag{17}$$

and

$$h''(z) = -3e^{3z} + \frac{2l}{n^{2/3}}e^{2z} + \frac{l}{n^{1/3}}e^z . \tag{18}$$

Let z_0 be the solution of $h'(z) = 0$. z_0 is located near $\frac{1}{3}\ln(a+1)$ because $a = (3l+1)/2$ is large. Note that z_0 can be obtained solving the cubic equation $h'(z) = 0$. Straightforward calculations leads to the following estimate

$$z_0 = \frac{1}{3}\ln\left(\frac{3}{2}(l+1)\right) + O\left(\frac{l^{1/3}}{n^{1/3}}\right) = \frac{1}{3}\ln(a+1) + O\left(\frac{l^{1/3}}{n^{1/3}}\right) \tag{19}$$

and

$$h(z_0) = \frac{a+1}{3}\ln(a+1) - \frac{a+1}{3} + O\left(\frac{l^{4/3}}{n^{1/3}}\right) . \tag{20}$$

We have also $h''(z_0) = -\left(3(a+1) + 2l/n^{1/3}e^{z_0} + l/n^{2/3}\right)$ and more generally

$$h^{(m)}(z_0) = -3^{m-1}(a+1) + A_m \frac{l}{n^{1/3}}e^{z_0} + B_m \frac{l}{n^{2/3}}e^{2z_0} . \tag{21}$$

Thus,

$$\int_{-\infty}^{+\infty} e^{h(z)}dz = e^{h(z_0)} \int_{-\infty}^{+\infty} \exp\left(h''(z_0)\frac{(z-z_0)^2}{2} + P(z-z_0)\right) dz \tag{22}$$

where P is a power series of the form $P(x) = (a+1)\sum_{i\geq 3} p_i x^i$ and $h''(z_0) < 0$. At this stage, one can consider $\exp\left(h''(z_0)\frac{(z-z_0)^2}{2}\right)$ as the main factor of the integrand. We refer here to the book of De Bruijn [4, §4.4 and §6.8] for more discussions about asymptotic estimates on integrals of the form given by (22) and we infer that

$$\int_{-\infty}^{+\infty} e^{h(z)} dz \sim \sqrt{-\frac{2\pi}{h''(z_0)}} \exp\left(h(z_0)\right) . \tag{23}$$

Using the Stirling formula for Gamma function, i.e., $\Gamma(t+1) \sim \sqrt{2\pi t}\frac{t^t}{e^t}$ and the fact that z_0 is located near $\frac{1}{3}\ln(a+1)$, $h(z_0) \sim \frac{(a+1)}{3}(\ln(a+1)-1)$ and $h''(z_0) \sim -3(a+1)$, we can see that (23) leads to (13) which is also the formula obtained by Janson in [6]. □

To estimate α_l, due to (2), it is convenient to compare the magnitudes of

$$\sum_{k=1}^{n} \frac{k^{(3l+1)/2}}{n^{l+1}} \exp\left(-\frac{k^3}{24n^2} + \frac{lk^2}{8n^2} + \frac{lk}{2n}\right) \tag{24}$$

and of

$$\sum_{k=1}^{n} \frac{k^4}{n^3} \frac{k^{(3l+1)/2}}{n^{l+1}} \exp\left(-\frac{k^3}{24n^2} + \frac{lk^2}{8n^2} + \frac{lk}{2n}\right) . \tag{25}$$

Also we need to compare (24) to the other "error terms" contained in the "big-ohs" of (2). Using the asymptotic value given by lemma 3, we easily obtain the estimates of the two quantities and we compute respectively $2^{(3l+3)/2}3^{(l-1)/2}\Gamma((l+1)/2)$ for (24) and $2^{(3l+11)/2}3^{l/2+5/6}\Gamma((l+1)/2+4/3)/n^{1/3}$ for (25). Thus, the term "$O(k^4/n^3)$" in (2) can be neglected if $l = o(n^{1/4})$ otherwise the quantity represented by (25) is not *small* compared to that represented by (24). Similarly, straightforward calculations using (3) show that the terms k/n, $1/k$, l^3/k^2, $l^{1/2}/k^{1/2}$ and $(l+1)^{1/16}/k^{9/50}$ can also be neglected. Using Stirling formula for Gamma function, lemmas 1 and 3, we have

$$\alpha_l \sim \frac{\rho_l}{2} 2^{\frac{3l+3}{2}} 3^{\frac{l-1}{2}} \Gamma(\frac{l+1}{2}) . \tag{26}$$

After nice cancellations, it results that:

Theorem 4 *In a randomly growing graph of n vertices, if $l, n \to \infty$ but $l = o(n^{1/4})$, the expected number of transitions $l \to l+1$, for all l-components, is $\alpha_l \sim 1$.* ◇

Note that in [7, p 301–306, §16–18], the authors already proved, by entirely different methods, that the most probable evolution of a random graph, when regarding the excess of connected component, is to pass directly from 1-component to 2-component, from 2-component to 3-component, and so on.

Similarly, as an immediate consequence of calculations above and [6, Theorem 9], we have:

Corollary 5 *As $n \to \infty$ and $l = o(n^{1/4})$, the expected number of vertices that ever belong to an l-component is $\mathbb{E}V_l \sim (12l)^{1/3} n^{2/3}$ and the expected order of the largest l-component that ever appears is $\mathbb{E}V_l^{max} = O(l^{1/3} n^{2/3})$.* $\diamond$

Note that these results answer the last remark in [6].

3 Conclusion

We briefly point out a remark concerning the restrictions on l in constrained graphs problems, i.e., the creation and growth of components with prefixed configurations. A possible way of investigations could be to search for similar results with those of Bender-Canfield-McKay [1,2] for supergraphs of a given graph H.

References

[1] Bender, E., Canfield, R. and McKay B. (1990). The asymptotic number of labelled connected graphs with a given number of vertices and edges. *Random Structures and Algorithms*, 1:127–169.

[2] Bender, E. A., Canfield, E. R. and McKay B. D. (1992). Asymptotic properties of labeled connected graphs. *Random Structures and Algorithms*, 3:183–202.

[3] Bollobas, B. (1985). *Random Graphs.* Academic Press, London.

[4] De Bruijn, N. G. (1981). *Asymptotic Methods in Analysis.* Dover, New-York.

[5] Janson, S. (1993). Multicyclic components in random graphs process. *Random Structures and Algorithms*, 4:71–84.

[6] Janson, S. (2000). Growth of components in random graphs. *Random Structures and Algorithms*, 17:343–356.

[7] Janson, S., Knuth, D. E., Luczak, T. and Pittel B. (1993). The birth of the giant component. *Random Structures and Algorithms*, 4:233–358.

[8] Janson, S., Luczak, T. and Rucinski A. (2000). *Random Graphs.* John Wiley, New York.

[9] Kolchin, V. F. (1999). *Random Graphs.* Encyclopedia of Mathematics and its Applications **53**. Cambridge University Press.

[10] Wright, E. M. (1980). The Number of Connected Sparsely Edged Graphs. III. Asymptotic results *Journal of Graph Theory*, 4:393–407.